

# A Simple Proof of the Mutual Incoherence Condition for Orthogonal Matching Pursuit

Jian Wang and Byonghyo Shim

Information System Laboratory

School of Information and Communication

Korea University, Seoul, Korea 136-713

Email: jwang@ipl.korea.ac.kr, bshim@korea.ac.kr

Phone: 82-2-3290-4842

## Abstract

This paper provides a simple proof of the mutual incoherence condition ( $\mu < \frac{1}{2K-1}$ ) under which  $K$ -sparse signal can be accurately reconstructed from a small number of linear measurements using the orthogonal matching pursuit (OMP) algorithm. Our proof, based on mathematical induction, is built on an observation that the general step of the OMP process is in essence same as the initial step since the residual is considered as a new measurement preserving the sparsity level of an input vector.

## Index Terms

Compressive sensing, orthogonal matching pursuit (OMP), restricted isometric property (RIP), mutual incoherence condition.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0012525) and the research grant from the second BK21 project.

# A Simple Proof of the Mutual Incoherence Condition for Orthogonal Matching Pursuit

## I. INTRODUCTION

As a sampling paradigm guaranteeing the reconstruction of sparse signal with sampling rate significantly lower than the Nyquist rate, compressive sensing (CS) has received considerable attention in recent years [1], [2], [3]. The main goal of the CS is to accurately reconstruct a high dimensional sparse vector using a small number linear measurements. Specifically, for a given matrix  $\Phi \in \mathbb{R}^{m \times n}$  ( $n > m$ ), the CS recovery algorithm generates an estimate of  $K$ -sparse vector  $\mathbf{x} \in \mathbb{R}^n$  from a set of linear measurements

$$\mathbf{y} = \Phi \mathbf{x}. \quad (1)$$

Although this task seems to be a severely ill-posed inverse problem, due to the prior knowledge of sparsity information,  $\mathbf{x}$  can be perfectly reconstructed via properly designed recovery algorithm. Among many greedy search algorithms developed for this purpose, OMP algorithm has received special attention due to its simplicity and competitive reconstruction performance [5].

Theoretical analysis of OMP to date has concentrated primarily on two fronts. The first approach is based on the restricted isometric property (RIP). A sensing matrix  $\Phi$  satisfies the RIP of order  $K$  if there exists a constant  $\delta$  such that [7]

$$(1 - \delta) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta) \|\mathbf{x}\|_2^2 \quad (2)$$

for any  $K$ -sparse vector  $\mathbf{x}$  ( $\|\mathbf{x}\|_0 \leq K$ ). In particular, the minimum of all constants  $\delta$  satisfying (2) is called the isometry constant  $\delta_K$ . Wakin and Davenport have shown that the OMP can reconstruct all  $K$ -sparse signals if  $\delta_{K+1} < \frac{1}{3\sqrt{K}}$  [2]. This result has been recently improved by Wang and Shim to  $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$  [6]. The second approach is based on the coherence parameter. The coherence parameter  $\mu$  of the sensing matrix  $\Phi$  is defined as

$$\mu = \max_{i \neq j} |\langle \varphi_i, \varphi_j \rangle|$$

where  $\varphi_i$  and  $\varphi_j$  are two column vectors of  $\Phi$ . When the columns of  $\Phi$  have unit norm and satisfy the mutual incoherence condition given by  $\mu < \frac{1}{2K-1}$ , the OMP will recover  $K$ -sparse

TABLE I  
OMP ALGORITHM

---



---

Input:	$\mathbf{y}, \Phi, K.$
Initialize:	$k = 0, \mathbf{r}^0 = \mathbf{y}, T^0 = \emptyset.$
While	$k < K$
	$k = k + 1.$
(Identify)	$t^k = \arg \max_j  \langle \mathbf{r}^{k-1}, \varphi_j \rangle .$
(Augment)	$T^k = T^{k-1} \cup \{t^k\}.$
(Estimate)	$\hat{\mathbf{x}}_{T^k} = \arg \min_{\mathbf{x}} \ \mathbf{y} - \Phi_{T^k} \mathbf{x}\ _2.$
(Update)	$\mathbf{r}^k = \mathbf{y} - \Phi_{T^k} \hat{\mathbf{x}}_{T^k}.$
End	
Output:	$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}: \text{supp}(\mathbf{x})=T^K} \ \mathbf{y} - \Phi \mathbf{x}\ _2.$

---

signal  $\mathbf{x}$  from the measurements  $\mathbf{y} = \Phi \mathbf{x}$  [5]. It is well known that this result is also applied to  $\ell_1$ -minimization approach [8].

In this work, we provide a simple proof of the mutual incoherence condition for the OMP using mathematical induction. Our proof is built on an observation that the general step of the OMP process is in essence same as the initial step since the residual is considered as a new measurement preserving the sparsity level of an input vector. The mutual incoherence condition for the OMP is formally described in the following theorem.

*Theorem 1* (Mutual incoherence condition for OMP): For any  $K$ -sparse vector  $\mathbf{x}$ , the OMP algorithm perfectly recovers  $\mathbf{x}$  from the measurements  $\mathbf{y} = \Phi \mathbf{x}$  if the coherence parameter  $\mu$  satisfies

$$\mu < \frac{1}{2K-1}. \quad (3)$$

## II. SIMPLE PROOF OF THEOREM 1

Before presenting the proof of Theorem 1, we provide lemmas useful in our analysis.

*Lemma 2 (Norm inequality [4]):* For  $\mathbf{A}$  and  $\mathbf{B}$  in  $\mathbb{R}^{m \times n}$ ,  $\alpha \in \mathbb{R}$ , and  $\mathbf{x} \in \mathbb{R}^n$ , following inequalities are satisfied:

$$\|\mathbf{A}\|_2 \leq \sqrt{mn} \|\mathbf{A}\|_{\max},$$

$$\|\mathbf{A} + \mathbf{B}\|_2 \leq \|\mathbf{A}\|_2 + \|\mathbf{B}\|_2,$$

$$\|\mathbf{AB}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{B}\|_2,$$

$$\|\alpha \mathbf{A}\|_2 = |\alpha| \|\mathbf{A}\|_2,$$

$$\|\mathbf{Ax}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x}\|_2,$$

where  $\|\mathbf{A}\|_2$  is the spectral norm of  $\mathbf{A}$  and  $\|\mathbf{A}\|_{\max}$  is the maximum absolute value of elements of  $\mathbf{A}$  (i.e.,  $\|\mathbf{A}\|_{\max} = \max_{i,j} |a_{i,j}|$ ).

*Lemma 3 (A direct consequence of RIP [7]):* Let  $I \subset \{1, 2, \dots, n\}$  and  $\Phi_I$  be the restriction of the columns of  $\Phi$  to a support set  $I$ . If  $\delta_{|I|} < 1$ , then for any  $\mathbf{u} \in \mathbb{R}^{|I|}$ ,

$$(1 - \delta_{|I|}) \|\mathbf{u}\|_2 \leq \|\Phi_I' \Phi_I \mathbf{u}\|_2 \leq (1 + \delta_{|I|}) \|\mathbf{u}\|_2.$$

*Lemma 4:* The isometric constant  $\delta_K$  for the sensing matrix  $\Phi$  satisfies

$$\delta_K \leq (K - 1) \mu.$$

*Proof:*

Using Lemma 2, we have

$$\begin{aligned} \|\Phi_T \mathbf{x}_T\|_2^2 &= \|\mathbf{x}_T' \Phi_T' \Phi_T \mathbf{x}_T\|_2 \\ &\leq \|\mathbf{x}_T'\|_2 \|\Phi_T' \Phi_T\|_2 \|\mathbf{x}_T\|_2 \\ &= \|\Phi_T' \Phi_T\|_2 \|\mathbf{x}_T\|_2^2 \end{aligned} \tag{4}$$

where  $T$  is the support of  $\mathbf{x}$  (a set with the locations of the non-zero elements of  $\mathbf{x}$ ),  $\mathbf{x}_T$  is a vector composed of the elements of  $\mathbf{x}$  indexed by  $T$ . Noting that  $(i, j)$ -th element of  $\Phi_T' \Phi_T$  is  $\langle \varphi_i, \varphi_j \rangle$  and  $\varphi_i$  is the unit norm vector, it is clear that

$$\Phi_T' \Phi_T = \begin{pmatrix} 1 & \langle \varphi_1, \varphi_2 \rangle & \cdots & \langle \varphi_1, \varphi_n \rangle \\ \langle \varphi_2, \varphi_1 \rangle & 1 & \cdots & \langle \varphi_2, \varphi_n \rangle \\ & & \cdots & \\ \langle \varphi_n, \varphi_1 \rangle & \langle \varphi_n, \varphi_2 \rangle & \cdots & 1 \end{pmatrix}. \tag{5}$$

Now, let

$$\Phi_T' \Phi_T = (1 - \mu) \mathbf{I} + \mathbf{A} \quad (6)$$

then

$$\mathbf{A} = \begin{pmatrix} \mu & \langle \varphi_1, \varphi_2 \rangle & \cdots & \langle \varphi_1, \varphi_K \rangle \\ \langle \varphi_2, \varphi_1 \rangle & \mu & \cdots & \langle \varphi_2, \varphi_K \rangle \\ & & \cdots & \\ \langle \varphi_K, \varphi_1 \rangle & \langle \varphi_K, \varphi_2 \rangle & \cdots & \mu \end{pmatrix} \quad (7)$$

and thus  $\|\mathbf{A}\|_{\max} = \mu$ . Hence,

$$\|\Phi_T' \Phi_T\|_2 = \|(1 - \mu) \mathbf{I} + \mathbf{A}\|_2 \quad (8)$$

$$\leq \|(1 - \mu) \mathbf{I}\|_2 + \|\mathbf{A}\|_2 \quad (9)$$

$$\leq 1 - \mu + \sqrt{K^2} \|\mathbf{A}\|_{\max} \quad (10)$$

$$= 1 + (K - 1) \mu \quad (11)$$

where (10) is from Lemma 2. Using (4) and (11), we have

$$\begin{aligned} \|\Phi_T \mathbf{x}_T\|_2^2 &\leq \|\Phi_T' \Phi_T\|_2 \|\mathbf{x}_T\|_2^2 \\ &\leq (1 + (K - 1) \mu) \|\mathbf{x}_T\|_2^2. \end{aligned}$$

Recalling the definition of the RIP that  $\delta_K$  is the minimum satisfying (2), we have

$$\delta_K \leq (K - 1) \mu.$$

■

### Proof of theorem 1

*Proof:* We will prove the theorem using induction. In the first iteration ( $k = 1$ ) of the OMP algorithm,  $t^k (= t^1)$  becomes the index of the column maximally correlated with the measurement  $\mathbf{y}$ , i.e.,

$$t^k = \arg \max_i |\langle \varphi_i, \mathbf{y} \rangle|. \quad (12)$$

Then, we have

$$|\langle \varphi_{t^k}, \mathbf{y} \rangle| = \max_i |\langle \varphi_i, \mathbf{y} \rangle| \quad (13)$$

$$\geq \frac{1}{\sqrt{K}} \|\Phi'_T \mathbf{y}\|_2 \quad (14)$$

$$\geq \frac{1}{\sqrt{K}} \|\Phi'_T \Phi_T \mathbf{x}_T\|_2 \quad (15)$$

$$\geq \frac{1}{\sqrt{K}} (1 - \delta_K) \|\mathbf{x}_T\|_2 \quad (16)$$

$$\geq \frac{1}{\sqrt{K}} (1 - (K - 1) \mu) \|\mathbf{x}_T\|_2 \quad (17)$$

where (15) is due to  $\mathbf{y} = \Phi_T \mathbf{x}_T$ , (16) and (17) follow from Lemma 3 and 4, respectively.

Now, suppose that  $t^k$  is not belonging to the support of  $\mathbf{x}$  (i.e.,  $t^k \notin T$ ), then

$$|\langle \varphi_{t^k}, \mathbf{y} \rangle| = \|\varphi'_{t^k} \Phi_T \mathbf{x}_T\|_2 \quad (18)$$

$$\leq \|\varphi'_{t^k} \Phi_T\|_2 \|\mathbf{x}_T\|_2 \quad (19)$$

$$= \sqrt{\sum_{i \in T} |\langle \varphi_{t^k}, \varphi_i \rangle|^2} \|\mathbf{x}_T\|_2 \quad (20)$$

$$\leq \sqrt{\sum_{i \in T} \mu^2} \|\mathbf{x}_T\|_2 \quad (21)$$

$$= \sqrt{K} \mu \|\mathbf{x}_T\|_2 \quad (22)$$

where (21) is from the definition of  $\mu$  ( $\mu = \max_{i \neq j} |\langle \varphi_i, \varphi_j \rangle|$ ). This case, however, will never occur if

$$\frac{1}{\sqrt{K}} (1 - (K - 1) \mu) \|\mathbf{x}_T\|_2 > \sqrt{K} \mu \|\mathbf{x}_T\|_2 \quad (23)$$

or

$$\mu < \frac{1}{2K - 1}. \quad (24)$$

In summary, if  $\mu < \frac{1}{2K - 1}$ , then  $t^k \in T$  for the first iteration of the OMP algorithm.

Now we assume that the former  $k$  iterations are successful ( $T^k = \{t^1, t^2, \dots, t^k\} \in T$ ) for  $1 \leq k \leq K - 1$ . Then it suffices to show that  $t^{k+1}$  is in  $T$  but not in  $T^k$  ( $t^{k+1} \in T \setminus T^k$ ). Recall from Table I that the residual at the  $k$ -th iteration of the OMP is

$$\mathbf{r}^k = \mathbf{y} - \Phi_{T^k} \hat{\mathbf{x}}_{T^k}. \quad (25)$$

Since  $\mathbf{y} = \Phi_T \mathbf{x}_T$  and  $\Phi_{T^k}$  is a submatrix of  $\Phi_T$ ,  $\mathbf{r}^k \in \text{span}(\Phi_T)$  and thus  $\mathbf{r}^k$  can be expressed as a linear combination of the  $|T| (= K)$  columns of  $\Phi_T$ . Accordingly, we can express  $\mathbf{r}^k$  as  $\mathbf{r}^k = \Phi \mathbf{x}'$  where the support (set of indices for nonzero elements) of  $\mathbf{x}'$  is contained in the support of  $\mathbf{x}$ . In this sense, it is natural to interpret that  $\mathbf{r}^k$  is a measurement of  $K$ -sparse signal  $\mathbf{x}'$  using the sensing matrix  $\Phi$ . Thus, if (24) is satisfied, we guarantee that  $t^{k+1} \in T$  at the  $(k+1)$ -th iteration. Noting that the residual  $\mathbf{r}^k$  is orthogonal to the columns already selected<sup>1</sup>, index of these columns is not selected again (see the identify step in Table I) and hence  $t^{k+1} \in T \setminus T^k$ . This concludes the proof.  $\blacksquare$

Thus far, we have shown that the OMP algorithm is working perfectly if the sensing matrix  $\Phi$  satisfies the condition  $\mu < \frac{1}{2K-1}$ . Interestingly, this condition is not only sufficient but also necessary. We prove this claim by showing that, even with slight relaxation of this condition ( $\mu = \frac{1}{2K-1}$ ), it is possible that the OMP algorithm cannot perfectly recover  $K$ -sparse signal. Note that our construction of  $\Phi$  is similar to Cai, Wang, and Xu's work for proving the tightness of mutual incoherence condition for  $\ell_1$ -minimization [9, Remark 3.2].

*Remark 1 (Necessity of  $\mu < \frac{1}{2K-1}$ ):* Suppose  $\Phi$  has normalized columns ( $\|\varphi_i\|_2 = \varphi_i' \varphi_i = 1$ ) and also satisfies  $\mu = \frac{1}{2K-1}$ . Then it is clear that  $\Phi' \Phi \in \mathbb{R}^{n \times n}$  has a unit diagonal and the absolute value of the off-diagonal elements is upper bounded by  $\frac{1}{2K-1}$ . Now consider

$$\Phi' \Phi = \begin{pmatrix} 1 & -\frac{1}{2K-1} & \cdots & -\frac{1}{2K-1} \\ -\frac{1}{2K-1} & 1 & \cdots & -\frac{1}{2K-1} \\ & & \cdots & \\ -\frac{1}{2K-1} & -\frac{1}{2K-1} & \cdots & 1 \end{pmatrix}.$$

Then  $\Phi' \Phi$  is symmetric and positive semi-definite matrix, and hence  $\Phi$  can be found by an eigen-decomposition of  $\Phi' \Phi$  [9]. Note that an  $n \times n$  matrix  $\mathbf{K}$  with  $(\mathbf{K})_{i,i} = a$  and  $(\mathbf{K})_{i,j} = b$  for  $i \neq j$  is invertible if and only if  $a + (n-1)b \neq 0$ . Hence, for  $a = 1$  and  $b = -\frac{1}{2K-1}$ ,  $\Phi' \Phi$  is not invertible for the choice of  $n = 2K$ . In this case, eigen-decomposition of  $\Phi' \Phi$  becomes

$$\Phi' \Phi = \mathbf{U} \mathbf{\Lambda} \mathbf{U}' \quad (26)$$

<sup>1</sup>Since  $\Phi_{T^k} \hat{\mathbf{x}}_{T^k}$  is a projection of  $\mathbf{y}$ , the error vector  $\mathbf{y} - \Phi_{T^k} \hat{\mathbf{x}}_{T^k}$  (which equals  $\mathbf{r}^k$ ) is orthogonal to the projection  $\Phi_{T^k} \hat{\mathbf{x}}_{T^k}$ .

and

$$\Phi = \sqrt{\Lambda} \mathbf{U}'$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ & & \cdots & & & & \\ 0 & 0 & \cdots & \lambda_l & \underbrace{0 \cdots 0}_{2K-l} \end{pmatrix}$$

and  $\{\lambda_i\}_{i=1,2,\dots,l}$  are  $l$  nonzero eigenvalues of  $\Phi' \Phi$ . Since the rank of  $\Phi$  is  $l (< 2K)$ , there exists a vector  $\mathbf{z} \in \mathbb{R}^{2K}$  (which by definition is  $2K$ -sparse vector) in the null space of  $\Phi$  obeying  $\Phi \mathbf{z} = \mathbf{0}$ . One can then divide  $\mathbf{z}$  into two  $K$ -sparse vectors  $\mathbf{x}_1$  and  $-\mathbf{x}_2$  (i.e.,  $\mathbf{z} = \mathbf{x}_1 - \mathbf{x}_2$ ). This gives  $\Phi \mathbf{x}_1 = \Phi \mathbf{x}_2$  so that the OMP algorithm fails to recover  $K$ -sparse vector. In fact, no reconstruction algorithm can always guarantee the perfect recovery of  $K$ -sparse vector under  $\mu = \frac{1}{2K-1}$ .

## REFERENCES

- [1] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4655–4666, Dec. 2007.
- [2] M. A. Davenport and M. B. Wakin, "Analysis of Orthogonal Matching Pursuit using the restricted isometry property," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4395–4401, Sept. 2010.
- [3] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathématique*, vol. 346, pp. 589–592, 2008.
- [4] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Univ Press, 1996.
- [5] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231–2242, Oct. 2004.
- [6] J. Wang and B. Shim, "New bounds for restricted isometric constants in orthogonal matching pursuit and orthogonal multi matching pursuit," preparing for submission to *IEEE Trans. Inf. Theory*. (Also, submitted to *Asilomar Conference* 2011)
- [7] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2230–2249, May 2009.
- [8] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2845–2862, Nov. 2001.
- [9] T. T. Cai, L. Wang, and G. Xu, "Stable recovery of sparse signals and an oracle inequality," *IEEE Trans. Inf. Theory*, vol. 56, pp. 3516–3522, July 2010.